Lepton-Flavor-Violating Decays Combined Analysis

Marj Corcoran, September 2007

1 Introduction

This note documents the combined LFV analyses for the decays $K_L \to \pi^0\mu e$, $K_L \to \pi^0\pi^0\mu e$, and $K_L \to \pi^0\pi^0\pi^0$ with $\pi^0 \to \mu e$. This last decay is exactly the same as $K_L \to \pi^0\pi^0\mu e$ but with a π^0 mass constraint added to the μe system.

This analysis combines the 97 and 99 data. With the advice of the godparent committee, I have decided to not worry about the fact that the box has been opened for some fraction of the data. I will use the likelihood method (described below) to define the signal region in all cases. The box was re-closed when we redefined the signal region.

In all three modes, the trigger used was Trigger 7, the e-mu trigger. The normalization modes come from trigger 2 $K_L \to \pi^+\pi^-\pi^0$ decays as normalization for $K_L \to \pi^0\mu e$, and trigger 1 $K_L \to \pi^0\pi^0\pi^0_D$ decays as normalization for $K_L \to \pi^0\pi^0\mu e$.

2 Common Analysis

In this section I describe the parts of the analysis that are common to all three modes. The numerical values of the cuts are given in the section specific to each mode, since there are some differences between modes. I have used version 6.0 of KTeVana throughout.

2.1 KTSPILL requirements

A KTSPILL cut is made on all spills, and bad spills are not used in Monte Carlo generation. The KTSPILL bits set are:

- Bit 1 = trigger problems
- Bit 2 = DMPT Ped exponent > 0
- Bit 3 = Bad DMPT capacitor
- Bit 4 = Blown QIE comparator
- Bit 5 = Dead DMPT
- Bit 8 = Broken dynode
- Bit 9 = Pipeline problems
- Bit 10 = Global CsI problem
- Bit 11 = ETOT problems
- Bit 12 = FERA ADC problems

- Bit 13 = Drift chamber problems
- Bit 14 = Veto problems
- Bit 16 = Muon problems
- Bit 17 = HCC problems
- Bit 21 = DAQ/L3
- Bit 22 = not 832/799 run
- Bit 28 = severe TRD problem
- Bit 29 = intensity spikes

.

2.2 General reconstruction

The reconstruction begins with tracking requirements. Exactly two tracks are required which must form a good two-track vertex within the fiducial decay volume. The upstream and downstream segments of the tracks are required to match in the magnet within 2mm. Both tracks are required to match to calorimeter clusters–a hardware cluster for the electron and a software cluster for the muon. One track is required to be consistent with an electron (E/p within 5% of 1 and 3x3 fusion $\chi^2 < 10$), and the other with a muon (less than 1 GeV of energy deposited in the calorimeter). The projection of the downstream segment of the muon track is required to match to hits in the muon counters in all three layers (MU2, MU3X, and MU3Y), to within a road determined by multiple scattering. The TRD is used to confirm particle ID, as described in detail in the sections describing each mode.

Along with the charged tracks, either two or four neutral clusters must be found. All neutral clusters are required to be hardware clusters and to have 3x3 fusion $\chi^2 < 10$.

Several cuts are made to remove events with accidental activity. The maximum energy in any of the SA or RC counters must be less than 0.3 GeV. Another effective anti-accidental cut is a limit on the maximum number of extra (not associated with a track) in-time hits in the drift chambers.

2.3 Back Anti cuts

To discriminate against events in which a photon goes down the beam hole, a cut on the BA1 energy is made for all modes. The BA is not perfectly modeled in the MC, so there is a bit of a data/MC mismatch. Table 1 shows that the data/MC ratio is similar for three two-track modes, so the use of a similar normalization mode should eliminate most of the error associated with the shortcomings of the BA modeling. The variation in the apparent flux due to varying the BA1 cut is included as part of the systematic error on the flux.

The BA configuration changed between the 97 and 99 data-taking periods, resulting in different distributions in this variable as shown in figure 1. Even so, I have cut on the same place (15 GeV) in both periods, since the major benefit comes from eliminating events in the overflow bin.

	BA< 10 data fraction	BA < 10 MC fraction	data/MC
K_{e3}	0.838±0.008	0.896±0.008	0.935±0.012
$K_{\mu 3}$	0.835±0.012	0.898±0.011	0.930±0.018
$K3\pi^0_D$	0.848±0.005	0.907±0.032	0.935±0.033

Table 1: Data/Monte Carlo comparison of the fraction of events for which the BA1 energy is less than 10 GeV. Although there is some data/MC mismatch, the data/MC ratio is stable over these two-track modes. These numbers are from a range of runs in the 99 data set.

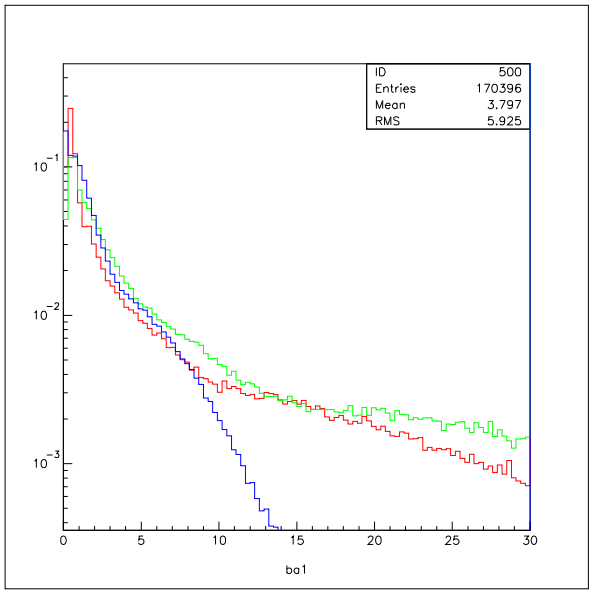

Figure 1: Distribution of BA1 energy for the three data-taking periods. The blue curve is 99 data, the red curve is winter97 and the green curve is summer97. The cut in this variable is set to 15 GeV in all cases

2.4 Likelihood Construction of the Signal Region

In all cases, the signal region is defined using the joint PDF constructed from the distributions in kaon mass and p_t^2. These distributions are similar enough in all modes that I use the same signal region for all modes. The mass distribution is fit with a Gaussian, and the p_t^2 distribution is fit with the sum of three exponentials. Figure 2 shows the distributions and the fits. These distributions are then normalized, and the product forms the joint PDF, which is the likelihood function. Figure 3 top shows the final likelihood function for $K_L \to \pi^0 \mu e$ signal Monte Carlo. In all cases a cut of $pdf > 10$ is made to define the signal region, which retains greater than 95% of the signal events after all other cuts are made. A cut of $pdf > 5$ defines the blind region. Figure 3 bottom shows how the new signal (red) and blind (blue) regions map onto the old signal box in the $p_t^2 - M_K$ plane.

For this procedure to be completely correct, the variables should be uncorrelated (although I have seen this method used when the variables are correlated, and it seems to work OK). Figure 4 shows the distribution in $(p_t^2 - < p_t^2 >) * (Kmass - < Kmass >)$. The mean of this distribution is the covariance, which is consistent

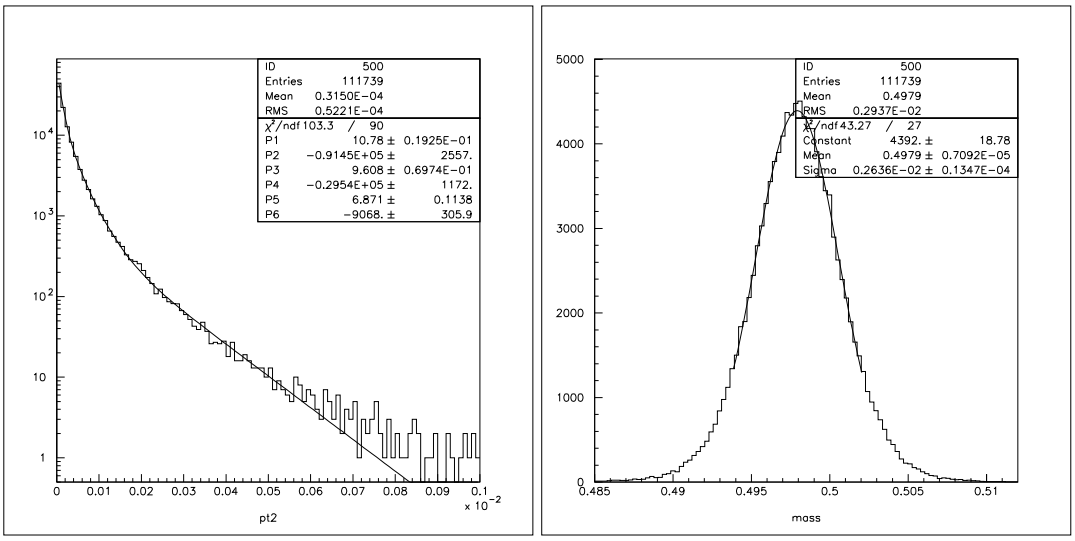

Figure 2: p_t^2 and Kaon mass fits used to define the likelihood function.

zero, indicating that these variables are indeed uncorrelated.

2.5 Cross-Trigger Issues

For all the LFV modes studied here, the trigger used is Trigger 7, which has a L1 requirement of at least one hit in MU3X and at least one hit in MU3Y. The normalization modes used (from trigger 2 or trigger 1) have no muon requirements, so there might be some trigger effects not compensated for by the normalization mode and not modeled in the Monte Carlo.

The muon counter efficiencies were carefully studied by Breese in his thesis. He studied the modeling of the cracks and the absolute efficiency, so that agreement between data and Monte Carlo was good to 0.5% or better. From Breese's work we can conclude that the muon banks are well-modeled in the Monte Carlo.

However as far as I could tell, Breese had not studied the actual L1 trigger efficiency, so I attempted to study the trigger efficiency using $K_{\mu 3}$ decays from trigger 2. To select a clean sample of $K_{\mu 3}$ decays, I require exactly two tracks, one of which is consistent with a muon in both the calorimeter and the TRDs. I also make cuts on the vertex χ^2, the track matching in the magnet, and energy in SA, RC, and BA1. I require the pion to have E/p between 0.2 and 0.8 and to be matched to a hardware cluster. Both the muon and pion are required to have momentum greater than 10 GeV. I also require the projection of the muon track to match with hits in all three of the muon planes, as determined by the routines MULAT2 (which uses the muon TDCs to determine which paddles have in-time hits) and MUMATCH2 (which matches the projection of the track with muon counter hits).

I have looked at both the L1 source TDCs and latches and find that the TDCs are slightly more efficient, but both show a L1 trigger efficiency of about 98%, which is lower than I expected. I look at all the TDC hits (the TDC handles up to 16, but I have never seen more then 4) and select as the "best hit" the one closest to the in-time peak at about -5ns. When all three muon banks have an in-time hit (as determined by the muon TDCs) matching the drift chamber track, the L1 TDC always has at least one hit, but a few percent of the time even the best hit is out of time. Figure 5 shows the MU3X and MU3Y TDC best-time distributions for events which pass the $K_{\mu 3}$ cuts.

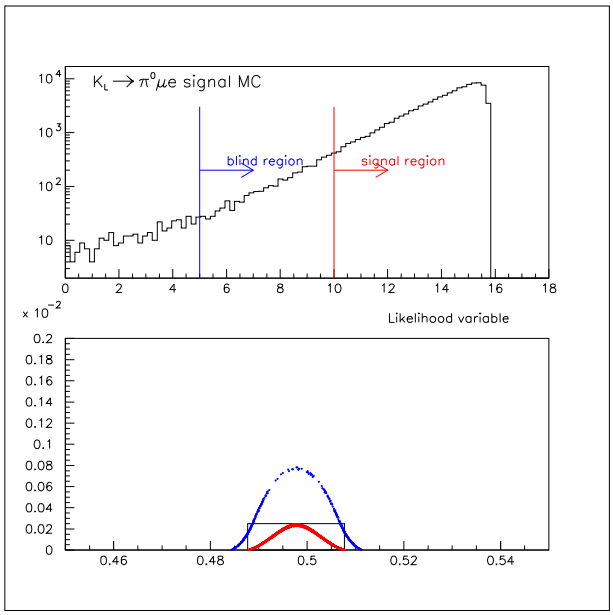

Figure 3: The top plot shows the distribution in the likelihood function *pdf* for $K_L \to \pi^0 \mu e$ signal Monte Carlo. The red arrow shows the signal region, and the blue arrow shows the blind region. The bottom plot is p_t^2 vs Kmass plane, showing the old signal region (box), the new signal region (red), and the blind region (blue).

I have tried several other cuts on the data to be sure I have a clean sample (pp0kine to exclude $K_L \to \pi^+\pi^-\pi^0$ and fiducial cuts around the beam holes), but I have not found any cuts that change this efficiency. I have to conclude this may be a true trigger inefficiency. If I make a timing cut of -20ns $\to$ +5ns in both MU3X and MU3Y, I find an efficiency of 97.2%. If I make a looser timing cut of -25ns $\to$ +5ns, I find an efficiency of 98.0%. I cannot rule out that the muon system is letting in out-of-time tracks (in spite of the timing cuts of -19ns $\to$ +4ns on the muon hits), or that the actual *trigger* (as opposed to the TDCs or latches) is more efficient. I will therefore take the L1MU3X/MU3Y trigger efficiency to be 98%, with a 2% systematic. The $K_{\mu 3}$ Monte Carlo does not exhibit any trigger inefficiency of this kind, so it is not corrected for by the Monte Carlo.

The other major difference between Trigger 7 and the normalization triggers is the number of HCC clusters required. Trigger 7 required one HCC cluster, and one track that was consistent with loose electron cuts at L3. Trigger 2 (normalization for $K_L \to \pi^0 \mu e$) had no HCC requirement, while Trigger 1 (normalization for $K_L \to \pi^0 \pi^0 \mu e$) required two tracks consistent with being an electron, and at least four HCC clusters. I can study the possible effects of any HCC trigger data/MC mismatch by varying the cut on the minimum cluster energy for the electron and photons in the normalization mode and including the change in apparent flux as part of the systematic error on the flux. The effect of this and other cut variations are discussed in the section on normalization mode and flux.

2.6 Combining 97 and 99 datasets and Including Systematics

For all modes studied here, the 97 and 99 datasets must be combined. The systematic errors also need to be taken into account.

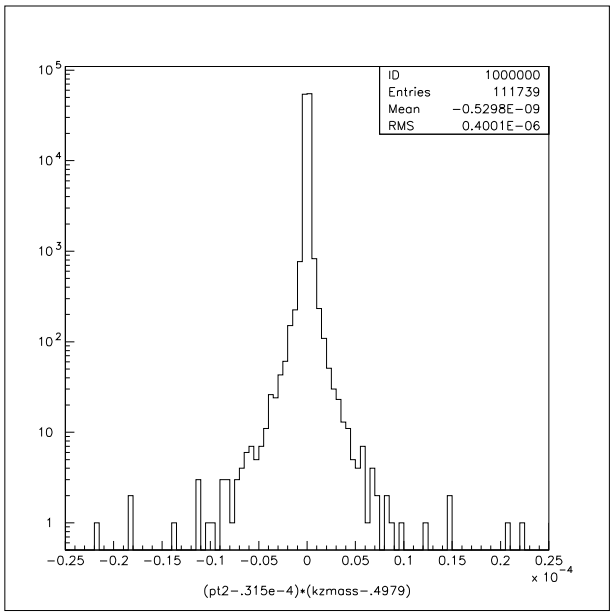

Figure 4: Distribution of $(p_t^2 - < p_t^2 >) * (M_K - < M_K >)$ from signal Monte Carlo. The mean of this distribution is the covariance, which is consistent with zero, indicating that these variables are uncorrelated.

To combine the 97 and 99 results, notice the total number of events observed would be $N_{tot} = BR(F_1 * \epsilon_1 + F_2 * \epsilon_2)$. Here F_1 and F_2 are the fluxes for the two datasets, and ϵ_1 and ϵ_2 are the acceptances. The quantity $(F_1 * \epsilon_1 + F_2 * \epsilon_2)^{-1}$ is the combined single event sensitivity (SES). So, for a given branching ratio, the mean number of expected signal events is given by N_{tot}. After discussion with the Godparent committee, we decided to use the Feldman-Cousins method to set the branching ratio limit. The main argument for using this method is that it is well-known and well-documented.

To determine a 90% CL limit, I add the observed number of events and the expected background (including errors) from the 99 and 97 data sets. I have constructed confidence bands using the Feldman-Cousins prescription, as described on more detail in section 6.

3 $K_L \to \pi^0\pi^0\mu e$

This mode has fairly low background due to the requirement of two neutral pions. The decay $\pi^0 \to \mu e$ can be searched for in the same data simply by putting a mass requirement on $M_{\mu e}$, the muon and electron invariant mass, after all cuts for $K_L \to \pi^0\pi^0\mu e$ have been made.

After the charged vertex is found, a search is made for exactly four neutral clusters forming two π^0s. All combinations of the clusters are tried, and the one for which the distance between the two neutral vertices is smallest is chosen as the candidate. The two vertices are required to be within three meters of each other for the event to be accepted. An average vertex is calculated, using the three vertices (one charged, two neutral) and their respective errors. The π^0 masses are recalculated using the average vertex, and a cut is made on these values as detailed below. An additional cut is made on the difference between the charged and each of the neutral vertices.

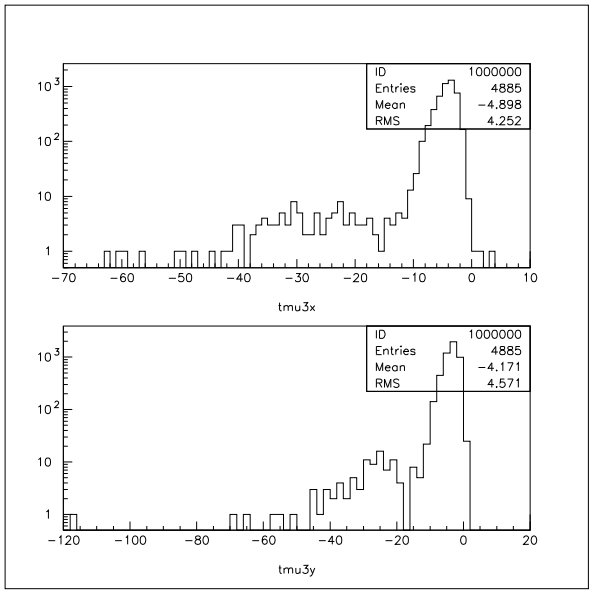

Figure 5: Time distributions for the best time in MU3X and MU3Y L1 trigger TDCs. This plot is for events which pass $K_{\mu 3}$ general cuts and for which MU2, MU3X and MU3Y all had in-time hits, as determined by the muon TDCs, matching the projection of the muon's drift chamber track.

3.1 TRD cuts

One source of background for these modes is $K_L \to \pi^0\pi^0\pi^0_D$ where an electron track is not well-measured in the calorimeter (for example because it is close to an edge) and at the same time is close enough to an accidental muon that it passes the muon-matching cut. For the $\pi^0 \to \mu e$ search, this is the dominant background. The TRDs are used in this analysis to discriminate against such events. I make a loose TRD cut to ensure that the track labeled as a muon by the calorimeter does not appear to be an electron as determined by the TRDs. The variable used is $prob_\mu$, the confidence level that the track in question is a muon or pion as determined by the TRDs. Electrons are therefore clustered near zero, so I require $prob_\mu$ >0.015, which retains 98.5% of muons but rejects 85% of electrons, as determined by $K_{\mu 3}$ and K_{e3} decays.

3.2 Kinematic cuts

Two kinematic variables, in the spirit of "ppi0kine" are used to help discriminate against backgrounds. I group together the two charged tracks and one of the neutral pions, and calculate the invariant mass of that system. From this information I find the square of the magnitude of the momentum of the other π^0 in the K rest frame. For an actual signal event, this quantity must be positive, but for many backgrounds this quantity is negative. There are two such variables ($kine1$ and $kine2$), one for each π^0. Figure 6 shows this kinematic variable for $K_L \to \pi^0\pi^0\mu e$ signal Monte Carlo and K_{e3} background Monte Carlo.

Another kinematic variable used in this analysis is the difference between the charged and neutral vertices, a cut which is similar to but not completely equivalent to a cut on the π^0 mass. Figure 7 shows this distribution for signal Monte Carlo. The cut is placed at ± 2.5 meters in this variable.

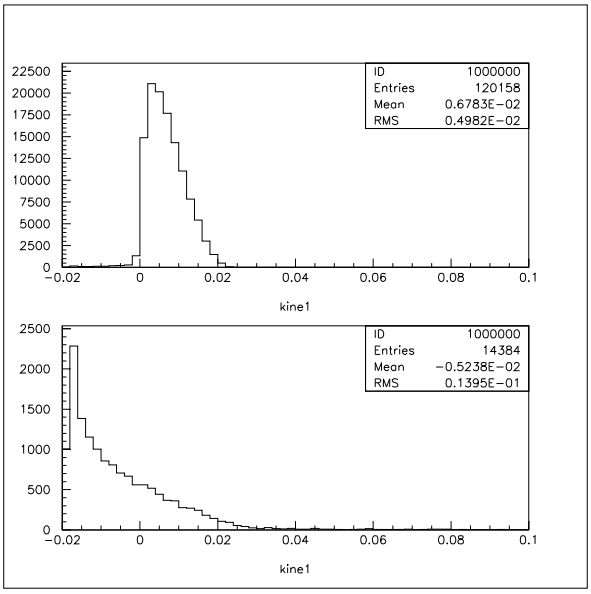

Figure 6: Kinematic variable equal to the square of the magnitude of one of the π^0's momentum in the K rest frame. The top plot is $K_L \to \pi^0\pi^0\mu e$ signal Monte Carlo, while the bottom plot is $K_{\mu 3}$ background Monte Carlo

.

3.3 Numerical values of cuts

- Track offsets in the magnet < 0.002m in all cases
- Z vertex location between 96 and 155 meters
- X and Y location of the vertex, projected to the calorimeter, is within the CsI beam holes
- 3x3 fusion $\chi^2 < 10$ for the electron and all neutral clusters
- Muon momentum > 8 GeV/c
- Energy in the calorimeter associated with the muon track < 1 GeV.
- Number of hardware clusters is exactly 5.
- Number of CsI clusters is < 13.
- Maximum energy deposited in any of the ring counters and spectrometer antis < 0.3 GeV
- Energy in BA1 < 15 GeV.
- Vertex χ^2 for the charged vertex < 20
- π^0 mass (calculated using the average vertex) between 0.132 and 0.138 GeV for both π^0
- Electron E/p between 0.95 and 1.05
- Projection of the downstream segment of the muon track matches to a hit in all three muon banks
- The difference between the charged and neutral vertices is less than 2.5 meters for both neutral pions.

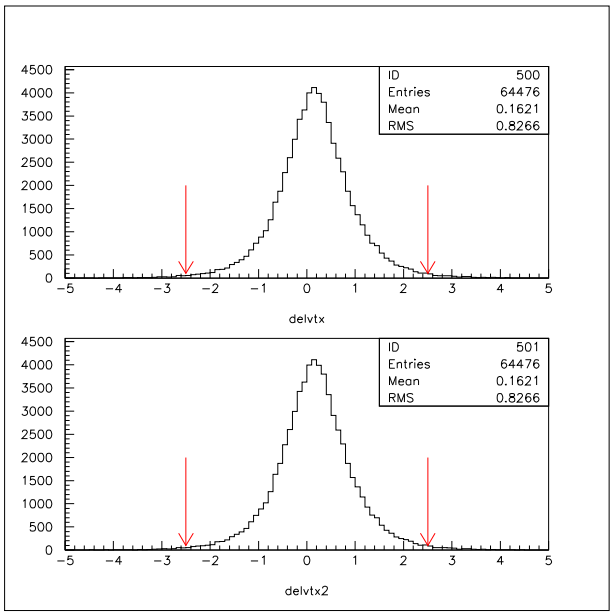

Figure 7: Difference in the charged and neutral vertices for signal Monte Carlo. The position of the cut in this variable is shown.

.

- Square of the magnitude of the π^0 momentum in the K rest frame ($kine1$ and $kine2$) is between 0 and 0.025 (GeV/c)^2 for each π^0.
- The number of extra in-time drift-chamber hit pairs is < 3 for both the upstream and downstream drift chambers
- The TRD information associated with the muon track is not consistent with an electron ($prob_\mu >$ 0.015).

The resulting signal acceptances for the three data periods are shown in Table 4. In all cases, the statistical error on the acceptance is 0.01%. These numbers include the 98.5% efficiency for the muon TRD cut and the 98% efficiency for the L1 trigger, both as determined from $K_{\mu 3}$ events from trigger 2.

3.4 Background studies

For $K_L \to \pi^0\pi^0\mu e$, the background appears to be a combination of all major modes. The next two figures compare search data with very loose cuts with a Monte Carlo sample (5% of a full flux) which combines (in the proper ratios) K_{e3}, $K_{\mu 3}$, $K_L \to \pi^0\pi^0\pi^0$, $K_L \to \pi^+\pi^-\pi^0$, and $K_L \to \pi^0\pi^0\pi^0_D$. Figure 8 shows $M_{\gamma\gamma}$, the two-photon mass, and $M_{\mu e}$, the invariant mass of the μe system. Figure 9 shows the charged vertex (top) and the electron E/p (bottom).

The combined background MC is a reasonable representation of search data at this level. But with even loose additional cuts, all MC events are eliminated. This sample is only 5% of a flux–generating a full flux would be very time-consuming. Moreover, trusting the MC to estimate the background to one part in 10^{10} is problematical. I therefore have tried to use the data itself as the best estimate of the background.

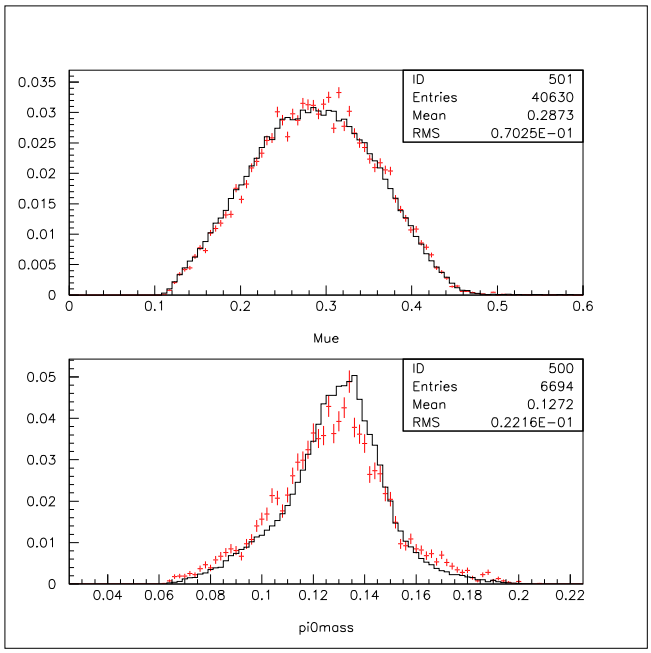

Figure 8: Comparison of the $M_{\mu e}$ mass (top) and the two-photon mass for 99 search data and the combined Monte Carlo sample, with the requirement that the charged and neutral vertices be within 20m of each other. The points are combined background Monte Carlo and the histogram is data.

I have tried to do this by using the number of events with the likelihood variable *pdf* in the region $-20 <$ *pdf* < 5 with various combinations of cuts. If all cuts are applied, there are not enough events to reliably extrapolate into the signal region. I have therefore defined three sets of cuts which should be reasonably independent, and I apply or remove them one set at a time. The cut sets are: kinematic cuts (kine1, kine2, $M_{\mu e}$, Δvertex, and M_{π^0}); anti-accidental cuts (BA1, SA/RC, DC upstream and downstream pairs, and number of CsI clusters); and particle ID cuts (all fusion χ^2 cuts, E/p for the electron, muon track matches in the muon counters, and TRD cut on the muon track). Figure 10 shows the distribution in the likelihood variable with all three of these cut sets have been removed. Table 2 summarizes the results of applying these cut sets one at a time, and with different combinations of two out of three. If the cut sets are really independent, the suppression factor obtained when applying both (such as Kinematic + ID) should equal the product of the suppression factors when each cut is applied individually (Kinematic * ID). As can be seen in the tables, this criteria is satisfied within the statistical errors, so we should be able to get the overall suppression factor as the product of the three individual factors. These numbers are for events with $-20 <$ *pdf* < 5.

We can now get an estimate for the background as follows. The overall suppression factor for these three cut sets is taken to be the product of the three individual factors. I take the distribution in *pdf* (with all cuts in the three sets removed) between -20 and 5, fit a straight line which I then extrapolate into the signal region of $10 <$ *pdf* < 15.8. The 1σ errors on the fit parameters are used to determine the systematic error on the background estimate in the signal region. For the 99 data, this procedure predicts 48.6 ± 10.5 events, and for the 97 data this procedure predicts 16.7 ± 9.4 events. Then applying the overall suppression factor, I obtain the background estimate in the signal region of 0.37 ± 0.11 events for the 99 data and 0.07 ± 0.05 events for the 97 data. The quoted errors include the statistical errors on the suppression factors as well as the systematic error arising from the $\pm 1\sigma$ variation of the fit parameters in the linear extrapolation into the signal region. I can use the same procedure to estimate the number of events in the blind region

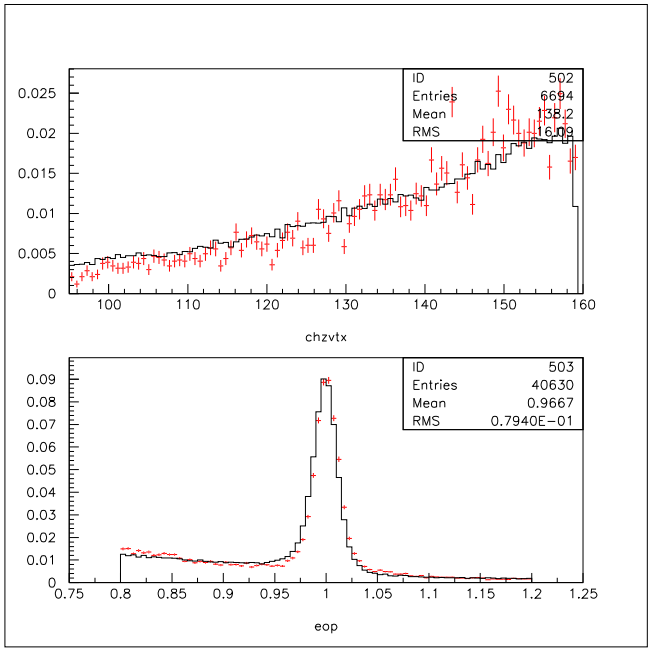

Figure 9: Comparison of the charged vertex location (top) and the electron E/p (bottom) , with the requirement that the charged and neutral vertices be within 20m of each other. The points are combined background Monte Carlo and the histogram is data.

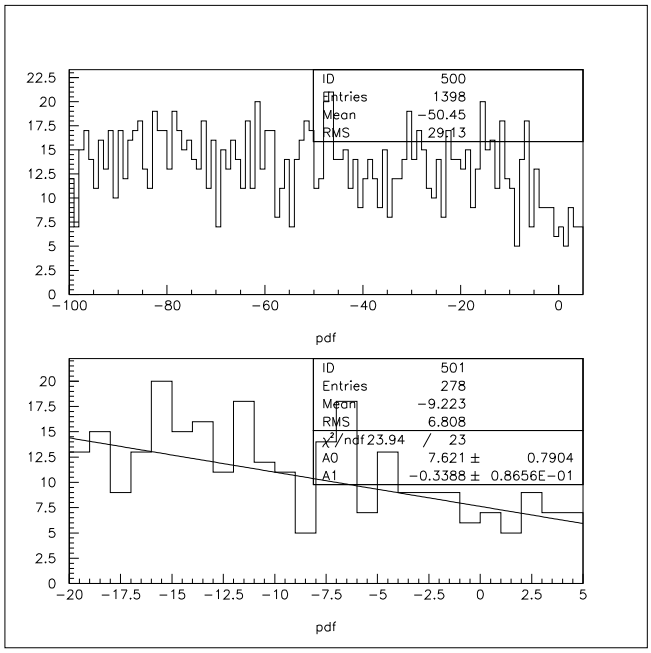

Figure 10: Distribution of the likelihood variable *pdf* with the three cuts sets as described in the text removed. The upper plot shows the range $-100 < pdf < 5$, and the lower plot shows an expanded scale $-20 < pdf < 5$ with a linear fit.

Cut set	Suppression factor
Kinematic	0.092 ± 0.016
Particle ID	0.273 ± 0.024
Anti-accidental	0.261 ± 0.024
Kinematic + ID	0.012 ± 0.009
Kinematic * ID	0.025 ± 0.005
Kinematic + Antiacc.	0.025 ± 0.009
Kinematic * Antiacc.	0.024 ± 0.005
ID + antiacc	0.077 ± 0.015
ID*antiacc	0.071 ± 0.009

Table 2: Effect of applying each of the three cut sets independently and applying them in pairs of two. The + symbol means both cuts have been applied to the data. The * symbol denotes the product of the two individual suppression factors. These numbers are for the 99 data and for $-20 < pdf < 5$.

$10 > pdf > 5$, which for the 99 data set is 0.34 ± 0.09 and for the 97 data set is 0.09 ± 0.04. Table 8 summarizes the background estimates for all three decay modes studied in this note.

3.5 Normalization Mode and Flux

The normalization mode for $K_L \to \pi^0\pi^0\mu e$ is $K_L \to \pi^0\pi^0\pi^0_D$ events taken from trigger 1. There are no muon or TRD cuts in the normalization sample, and there is a cut on M_{ee} >0.015 GeV to ensure that the tracks are far enough apart to be well-reconstructed. Otherwise, all the cuts are identical to the signal mode.

The following figures compare normalization data from the 1997 data-taking period to the Monte Carlo. Figure 11 shows the K mass; figure 12 shows the π^0 mass; figure 13 compares the electron E/p; figure 14 shows the charged vertex; figure 15 shows the invariant mass of the two electrons. There is a clear shift between data and Monte Carlo in the electron E/p and to some extent the Kaon mass. This shift is also present in the 99 data, but to a lesser extent. I have tried to deal with the data/Monte Carlo mismatch by keeping the cut on E/p fairly loose, and I include the change in apparent flux as this cut is varied as part of the systematic error.

Table 4 shows the normalization mode acceptance and apparent flux for each of the data periods.

3.6 Systematics

I have studied the systematics on the flux measurement by varying the cuts to see the effect on the flux. All cuts listed in section 3.3 were varied, except those applied to the muon track, since the normalization mode has no muon track. The cut that had the single largest effect was the BA1 cut; removing it completely has about a 4% effect on the 99 flux, but it was less than a 2% effect for both periods in 97. The second-largest effect was due to the cut on extra in-time pairs in the drift chambers. The default allows at most two extra hits in either the upstream or downstream chambers. Changing this from 2 to 1 lowered the apparent flux by 1.5-2% in all cases. Only variations larger than 1% in at least one of the data periods were included in the total systematic. Table 3 lists the cut variations included in the flux systematic. The total systematic

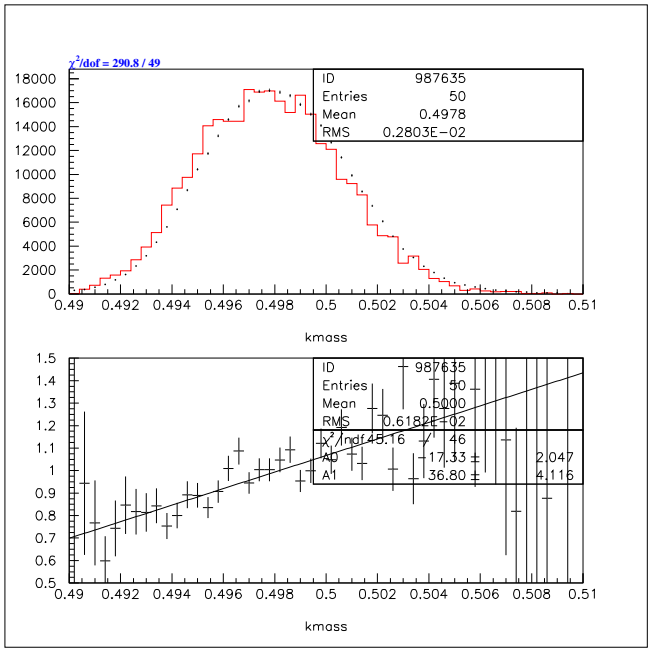

Figure 11: MC vs. data for the normalization mode: Kaon Mass. The red histogram is the MC and the black points are the data. The second plot is data/MC.

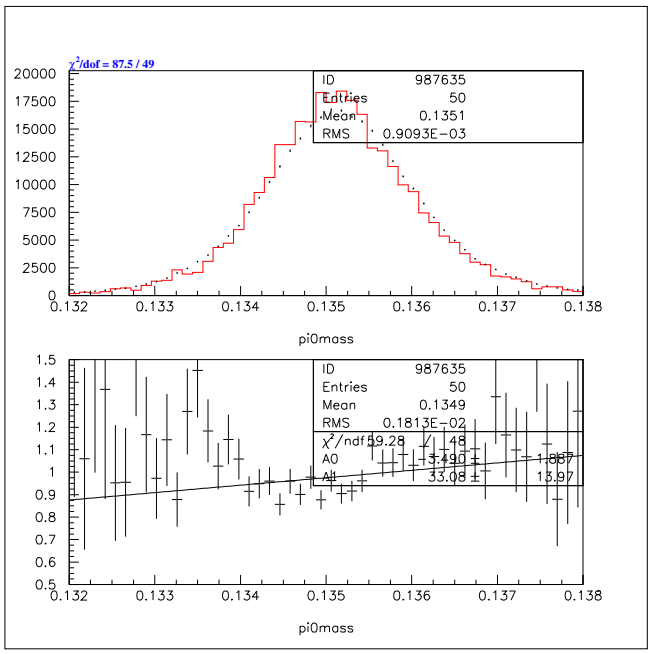

Figure 12: MC vs. data for the normalization mode: π^0 Mass. The red histogram is the MC and the black points are the data. The second plot is data/MC.

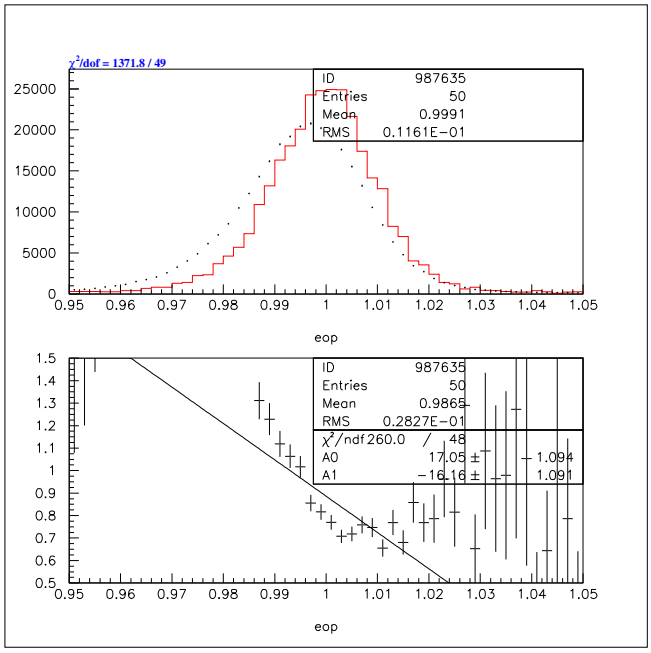

Figure 13: MC vs. data for the normalization mode: electron E/p. The red histogram is the MC and the black points are the data. The second plot is data/MC.

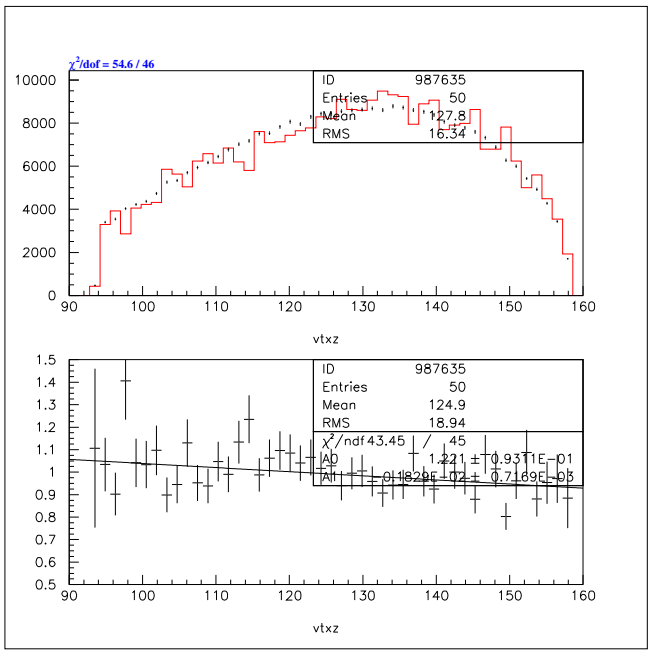

Figure 14: MC vs. data for the normalization mode: Charged z vertex. The red histogram is the MC and the black points are the data. The second plot is data/MC.

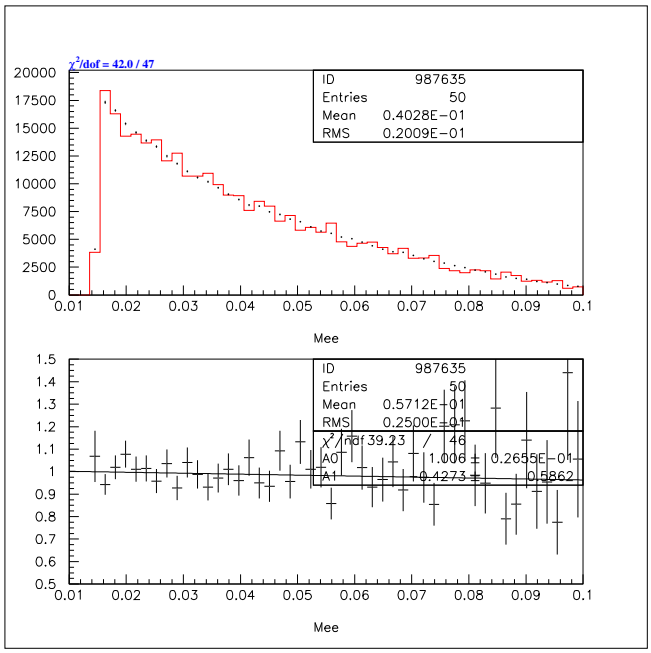

Figure 15: MC vs. data for the normalization mode: M_{ee}. The red histogram is the MC and the black points are the data. The second plot is data/MC.

error was taken as the sum in quadrature of these contributions.

The other systematic (discussed in section 2.5) is due to the uncertainty in the L1 trigger efficiency. I determined an inefficiency of 2%, but I take the systematic on this number to be 100% of itself.

3.7 $\pi^0 \to \mu e$

The analysis for this decay is exactly the same as for $K_L \to \pi^0\pi^0\mu e$, with the addition of a cut on $M_{\mu e}$, the $\mu - e$ invariant mass. The results for signal mode acceptance are shown in table 4. The flux of π^0s is determined from number of K_L decays and the $K_L \to \pi^0\pi^0\pi^0$ branching ratio. The flux of π^0s for all three data periods is also shown in table 4.

The signal region in the μe invariant mass is $0.132 < M_{\mu e} < 0.138 GeV/c^2$, but even a looser cut of $M_{\mu e} <$ 0.015 GeV/c^2 eliminates most of the background, as can be seen in the top plot of figure 8. Figure 16 (top) shows the distribution for $M_{\mu e}$ for $K_L \to \pi^0\pi^0\pi^0_D$ Monte Carlo with no cuts other than $0.12 GeV/c^2 < M_{\mu e} < 0.15 GeV/c^2$. The distribution is consistent with being flat in this region. If I apply all cuts except the muon-track matching cut and TRD cut, the $M_{\mu e}$ distribution for $K_L \to \pi^0\pi^0\pi^0_D$ Monte Carlo is as shown in the middle plot of figure 16. The bottom plot shows the 99 search data with identical cuts to the middle plot. The Monte Carlo is 3x a full flux, so the number of events in the data is consistent with being completely due to $K_L \to \pi^0\pi^0\pi^0_D$ background.

To get an estimate of the background for this mode, I use the data as was done for $K_L \to \pi^0\pi^0\mu e$. If I remove the three cut sets described in section 3.4 and require $0.12 GeV/c^2 < M_{\mu e} < 0.15 GeV/c^2$ (5x the size of the signal region) and $-20 < pdf < 5$, 46 events remain in the 99 data sample. I fit the distribution to a straight line and extrapolate into the signal region, using $\pm 1\sigma$ on the fit parameters as the systematic. I then apply the suppression factors for each cut set as determined from the data, and apply an additional

Cut variation	Variation 99 data	Variation summer 97	Variation winter 97
BA1 removed	3.7%	1.5%	0.06%
E/p (.96 to 1.04)	0.9%	1.2%	1.0%
Extra DC hits	2.0%	2.0%	1.4%
Cluster energies > 3 GeV	2.0%	1.0%	1.2%
Track matching in magnet	2.3%	1.0%	1.2%
Neutral π^0 mass	0.8%	1.2%	0.7%
Charged π^0 mass	0.8%	1.5%	1.0%
Fusion χ^2	0.8%	1.1%	1.2 %
Total	5.5%	3.8%	3.0%

Table 3: Summary of systematic errors on the apparent flux due to cut variations for the $K_L \to \pi^0\pi^0\pi^0_D$ normalization mode

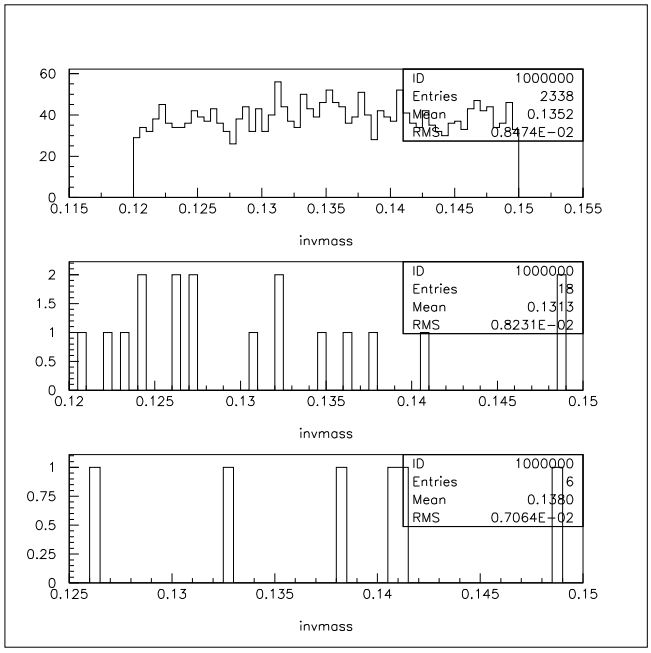

Figure 16: Distribution for $M_{\mu e}$. The top plot shows $K_L \to \pi^0\pi^0\pi^0_D$ Monte Carlo (3x full flux), with only the requirement $0.12 GeV/c^2 < M_{\mu e} < 0.15 GeV/c^2$ (five times the range of the signal region). The middle plot is Monte Carlo with all cuts in place except the muon ID cuts (muon-track matching and TRD). The bottom plot is 99 search data with identical cuts as the middle plot.

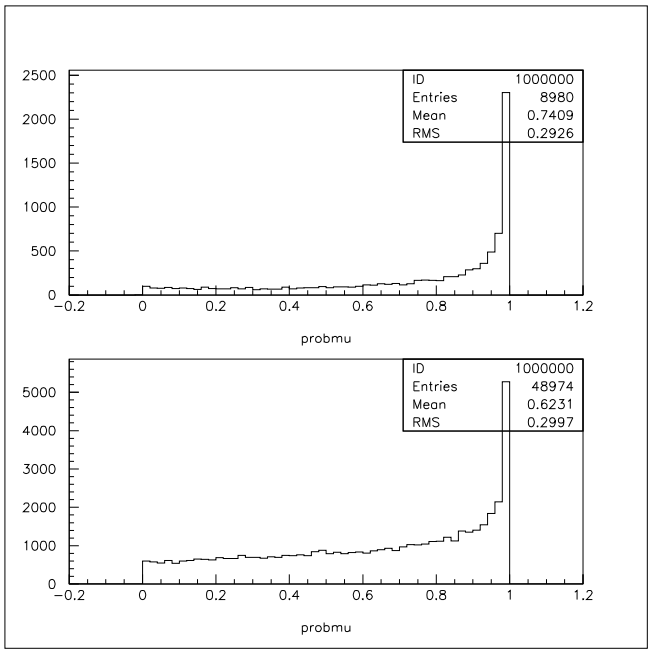

Figure 17: TRD pion probability for the track labeled muon. The top plot is $K_{\mu 3}$ Monte Carlo and the bottom plot is $K_{\mu 3}$ data from trigger 2. A loose anti-electron cut of $0.015 < prob_\mu$ is used.

factor of 1/5 for the range in $M_{\mu e}$, to get a background estimate of 0.015 ± 0.011 for the 99 dataset.

I can use the $K_L \to \pi^0\pi^0\pi^0_D$ Monte Carlo to estimate the background as a cross-check, since, with the $M_{\mu e}$ cut, the background is consistent with being completely from $K_L \to \pi^0\pi^0\pi^0_D$. If I apply all cuts no events are left in the signal region, so I use "cut set/suppression factor" method as was used with the data to arrive at a background estimate of 0.023 ± 0.013, consistent with the previous background estimate.

In determining the background estimate from $K_L \to \pi^0\pi^0\pi^0_D$ Monte Carlo, I have applied the TRD cut on the muon track to the Monte Carlo events. One could question whether it is valid to apply TRD cuts to the Monte Carlo, but figure 17 shows that it is reasonable. This figure shows the distribution in the TRD variable $prob_\mu$, the CL that the track labeled as the muon looks like a pion to the TRDs. The top plot is $K_{\mu 3}$ Monte Carlo, and the bottom plot is $K_{\mu 3}$ events from trigger 2 data. I make a cut at $prob_\mu > 0.015$ to eliminate events in which the track labeled as the muon looks very electron-like in the TRDs. The Monte Carlo should make a reasonable estimate of this rejection.

The background estimate for the 97 data is obtained in the same way from the data itself. Since the overall rate of background is lower in 97, I used a range of $-50 < pdf < 5$ to get enough events for a reasonable extrapolation into the signal region. The resulting background estimate is 0.013 ± 0.01 events. The background estimate from a 2x flux $K_L \to \pi^0\pi^0\pi^0_D$ Monte Carlo is 0.01 ± 0.007.

Table 4 summarizes the signal acceptance, normalization acceptance and flux for $K_L \to \pi^0\pi^0\mu e$ and $\pi^0 \to \mu e$ modes. The background estimates are summarized in table 8 for all modes.

Mode and data period	**Signal Acc.**	**Norm. acc.**	**Flux** (10^{11})
$K_L \to \pi^0\pi^0\mu e$, 99	2.04%	0.242%	3.60 $\pm$0.22
$K_L \to \pi^0\pi^0\mu e$ summer97	1.87%	0.195%	1.13 $\pm$0.054
$K_L \to \pi^0\pi^0\mu e$ winter97	2.00%	0.207%	1.63 $\pm$0.068
Total Flux of K_Ls			6.36 $\pm$0.240
$\pi^0 \to \mu e$ 99	1.73%	0.242%	2.28 $\pm$0.125
$\pi^0 \to \mu e$ summer97	1.53%	0.195%	0.72 $\pm$0.027
$\pi^0 \to \mu e$ winter97	1.65%	0.207%	1.03 $\pm$0.031
Total Flux of π^0s			4.03 $\pm$0.132

Table 4: Final results for acceptances and fluxes for the decay modes $K_L \to \pi^0\pi^0\mu e$ and $\pi^0 \to \mu e$.

4 $K_L \to \pi^0\mu e$

The reconstruction in this mode is the same as $K_L \to \pi^0\pi^0\mu e$, except that only two neutral clusters are required, and they must reconstruct to a π^0 mass using the charged vertex. The electron and both photons must be hardware clusters, the muon must be matched to a software cluster, and no extra hardware clusters can be present.

This mode has significantly more background since it requires only one π^0. Attempts to reduce this background led to some additional cuts not included in the $K_L \to \pi^0\pi^0\mu e$ analysis, as well as harder cuts for some variables. The dominant background is K_{e3} decays with two accidental photons and a pion decay or punch through. The K_{e3} background results in a $M_{\gamma\gamma}$ distribution that has a linear shape and no hint of a π^0 peak. One way to reduce this background is to cut as hard as possible on $M_{\gamma\gamma}$. The location of the cut on $M_{\gamma\gamma}$ was chosen by maximizing $S/\sqrt{B}$, where S is the signal acceptance and B is the expected background. The optimal cut was found to be at $\pm 1.4\sigma$ about the π^0 mass.

Since the K_{e3} background involved accidental activity, Angela looked for additional anti-accidental cuts. One which was found to be effective is the number of upstream track segments in both the x and y views of the drift chambers. These variable are called *nxtrks* and *nytrks*, and a cut is placed on each as detailed in the list below. The cut on extra in-time drift chamber hits is significantly harder in this analysis than in the $K_L \to \pi^0\pi^0\mu e$ analysis.

In order to ensure that $K_L \to \pi^+\pi^-\pi^0$ events are eliminated, a loose TRD cut is placed on the electron track so that a pion does not fake an electron. The value of the cut is different for 99 and 97, due to the different performances of the TRDs in these periods. See figure 3.25 in Angela's thesis.

Finally, there are three kinematic variables which help reduce backgrounds. The traditional *pp*0*kine* discriminates against $K_L \to \pi^+\pi^-\pi^0$ background. If the decay is assumed to be a K_{e4} decay, we can calculate the momentum of the unseen neutrino (called *ke*4*kine*). If the decay really is a K_{e4}, this quantity will be positive, while for signal Monte Carlo this quantity is usually negative. A cut on *ke*4*kine* provides powerful discrimination against K_{e4} decays. One final kinematic variable is the square of the longitudinal momentum of the π^0, calculated in the K rest frame, assuming the decay is the signal mode. A true signal decay would have this quantity positive, so we can eliminate some background by requiring this variable to be positive. Figures 3.40, 3.41, and 3.42 in Angela's thesis show these kinematic variables.

4.1 Numerical values of cuts

The numerical values of cuts used in this analysis are given below.

- KTSPILL cuts same as $K_L \to \pi^0\pi^0\mu e$ analysis
- Track offsets in the magnet < 0.002m in all cases
- Z vertex location between 96 and 155 meters
- X and Y location of the vertex, projected to the calorimeter, is within the CsI beam holes
- 3x3 fusion $\chi^2 < 10$ for the electron and all neutral clusters
- Muon momentum > 8 GeV/c
- Energy in the calorimeter associated with the muon track < 1 GeV.
- Maximum energy deposited in any of the ring counters and spectrometer antis <0.3 GeV
- Energy in BA1 < 15 GeV.
- Vertex χ^2 for the charged vertex < 20
- π^0 mass within $\pm 1.4\sigma$ of the mean.
- Electron E/p between 0.95 and 1.05
- Projection of the downstream segment of the muon track must match to a hit in all three muon scintillator banks.
- Square of the magnitude of the π^0 momentum in the K rest frame (assuming the signal decay mode) is between 0 and 0.025 (GeV/c)^2.
- The number of extra in-time drift-chamber hit pairs is < 3 for the downstream drift chambers and 0 for the upstream chambers. (Note that this cut is harder than for $K_L \to \pi^0\pi^0\mu e$)
- The TRD information associated with the electron track is consistent with being an electron with 98% efficiency. (Prob_e < 0.192 for the 99 dataset and < 0.04 for the 97 dataset).
- Total number of CsI clusters < 13.
- The number of upstream track candidates in x is less than 10 and in y is less than 5.
- Number of hardware clusters is exactly three.
- Assuming a K_{e4} decay, the momentum of the unseen neutrino in the K rest frame (ke4kine) satisfies $-0.08 < ke4kine < -0.005$.
- Assuming K3pi decay, require pp0kine <-0.06

Signal acceptances for all data periods are shown in Table 7. These numbers include the muon trigger efficiency of 98% as discussed in section 2.5 as well as the TRD electron efficiency of 98%.

4.2 Backgrounds

Angela studied many sources of background. She found that double decays, hyperon decays, and $K_{\mu 3}$ decays did not contribute significantly to the background in this mode. Other possible sources of background are discussed below.

4.2.1 K_{e4} background

One source of background is K_{e4} decays with a low energy neutrino, which is estimated from Monte Carlo. The original estimate of background due to this source was flawed due to the pi-mu decay bug uncovered by Rick Kessler, inadequate modeling of pion punch through, and incorrect form factors in the Monte Carlo. All of these problems of these have been corrected, including using the most recent NA48 form factors for Monte Carlo generation. Using a 40x flux Monte Carlo sample of K_{e4} decays, I find background from this source in the new signal region ($pdf > 10$) to be 0.075 ± 0.043 for the 99 data set and 0.025 ± 0.025 for the 97 data set. The blind region has more K_{e4} events, with 1.20 ± 0.17 for 99 and 0.45 ± 0.11 for 97. This background is larger than the previous estimate, but it is still small in the signal region.

4.2.2 $K_L \to \pi^+\pi^-\pi^0$ Background

$K_L \to \pi^+\pi^-\pi^0$ events can fake the signal if one charged pion decays and the other fakes an electron. To help eliminate this source of background, the TRD information on the electron track is used to ensure a pion does not fake an electron. These events are further suppressed by the fact that they reconstruct very low in kaon mass since one pion has been assigned an electron mass and the other has been assigned a muon mass. To confirm that this background is negligible, I have estimated the $K_L \to \pi^+\pi^-\pi^0$ background from Monte Carlo, but without the electron ID cuts in the calorimeter or TRD. A MC sample of 500M $K_L \to \pi^+\pi^-\pi^0$ events, and no electron ID cuts of any kind, yields no events in the signal region or even near the signal region. Figure 18 shows the p_t^2 vs kaon mass for this Monte Carlo sample. Based on K_{e3} events from Trigger 2 in 99 data, I find that the π/e rejection factor for the from E/p and fusion χ^2 cuts to easily larger than 600. The rejection factor from the loose TRD cut made here is about 5. A sample of 500M generated $K_L \to \pi^+\pi^-\pi^0$ events with no electron-ID cuts then translates into 1.5×10^{12} $K_L \to \pi^+\pi^-\pi^0$ decays with the electron cuts in place. This sample then corresponds to about 35x the actual number of $K_L \to \pi^+\pi^-\pi^0$ decays in the 99 data. So background from this source is neglible.

4.2.3 K_{e3} Background

After all cuts we believe the dominant background remaining is due to K_{e3} decays with two accidental photons faking a π^0. The $M_{\gamma\gamma}$ distribution for this background shows no hint of a π^0 peak, as shown in figure 19. This figure shows the $M_{\gamma\gamma}$ distribution in the search data with all cuts applied except the $M_{\gamma\gamma}$ cut.

Anglea did an extensive Monte Carlo of the K_{e3} background. The flat distribution in $M_{\gamma\gamma}$ as well as p_t^2 is well-reproduced in the Monte Carlo. (see Appendix C in Angela's thesis). But we have always favored using the data itself to estimate this background for two reasons: the Monte Carlo is not reliable at the level of one part in 10^{10}, and the method described below will include all sources of background which do not peak in $M_{\gamma\gamma}$.

To discriminate against this background we should cut as hard as possible on the π^0 mass in the $M_{\gamma\gamma}$ distribution. But a hard cut also impacts the signal acceptance significantly. The best procedure is to

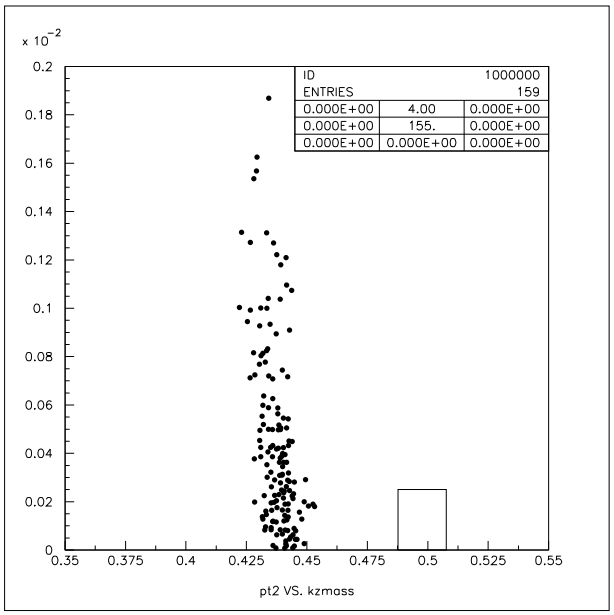

Figure 18: P_t^2 vs. Kmass plot for 500M K3pi Monte Carlo events. All cuts except electron ID cuts have been applied. The old signal region is shown.

optimize the cut by maximixing a ratio such as $S/\sqrt{B}$ where S is the signal accepance and B is the expected number of background events. Maximizing this quantity leads me to cut at $\pm 1.4\sigma$ around the π^0 mass. This is somewhat looser than the $\pm 1.25\sigma$ in Angela's analysis.

I have estimated the background in this mode in two ways: by an extrapolation in *pdf*, as was done for $K_L \to \pi^0\pi^0\mu e$, and by using the $M_{\gamma\gamma}$ sidebands and counting events in the signal and blind regions in *pdf* which pass all other cuts.

The extrapolation in *pdf* used all events with $M_{\gamma\gamma}$ in the range of 0.11 GeV to 0.16 GeV. I looked at the distribution in *pdf* from -20 to 5, fit a straight line, and extrapolated into the blind and signal regions. As was done before, I use the $\pm 1\sigma$ errors on the fit parameters to estimate the systematic error. Then I corrected for the width of the $M_{\gamma\gamma}$ region–this estimate uses a $M_{\gamma\gamma}$ range of 0.11 to 0.16 GeV, while the final signal region is $\pm 0.0040 GeV$ about the π^0 mass (corresponding to the optimized cut of $\pm 1.4\sigma$).

Another way to estimate the K_{e3} background is to use the $M_{\gamma\gamma}$ sidebands. If I remove the central region in $M_{\gamma\gamma}$ (0.132 - 0.138 GeV) but apply all other cuts, I and can then look in the signal and blind regions. The number of events found can then be scaled by the relative ranges in $M_{\gamma\gamma}$. These background estimates for both the signal region ($pdf > 10$) and the blind region ($5 < pdf < 10$) are shown in Table 5. The background estimates from the two methods agree, but the errors are smaller for the sideband determination, so I will use that. The sideband method has the advantage of specifically excluding any K_{e4} contribution, so that the K_{e4} contribution determined from Monte Carlo can be added to the K_{e3} background.

5 Flux and Acceptance

The normalization mode for this decay is $K_L \to \pi^+\pi^-\pi^0$, taken from Trigger 2. Chapter 3 in in Angela's thesis shows many comparisons between normalization data and Monte Carlo. Table 6 gives results on cut

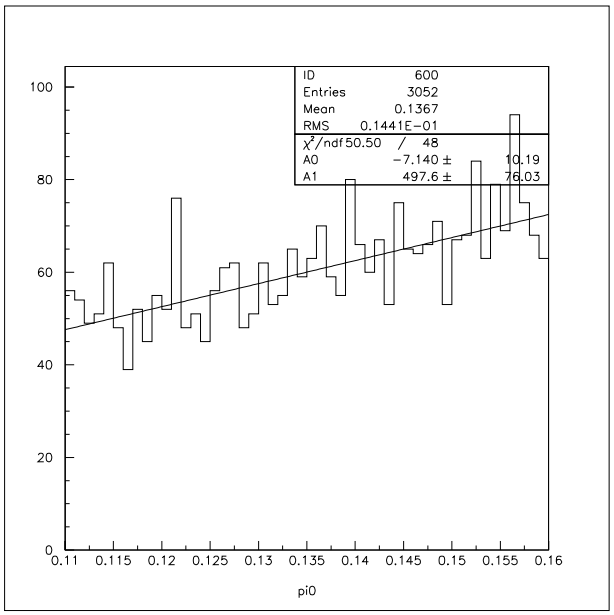

Figure 19: $M_{\gamma\gamma}$ distribution for the 99 data, with all cuts in place except the $M_{\gamma\gamma}$ cut.

Method	99 signal	99 blind	97 signal	97 blind
pdf extrap.	0.55 ±0.54	0.98 ± 0.39	0.43 ± 0.43	0.57 ± 0.29
$M_{\gamma\gamma}$ sidebands	0.28 ± 0.16	1.83 ± 0.41	0.28 ± 0.16	0.73 ± 0.26

Table 5: Summary of backgrounds estimated from the data for $K_L \to \pi^0\mu e$. The signal region is $pdf > 10$, and the blind region is $5 < pdf < 10$.

variations. All of the final cut variables in section 4.1 which were applicable were varied. Cuts involving electrons could not be varied, since the normalization mode has no electron. Only those cuts for which the change in apparent flux was at least 1% in at least one data period are shown in this table.

The systematic errors on the flux are somewhat larger for $K_L \to \pi^0\mu e$ than for $K_L \to \pi^0\pi^0\mu e$, due mainly to data/MC mismatch in some of the cuts on accidental activity and the hard cuts madeon these variables. But I have found that relaxing these cuts definitely lets in more background, so we are better off keeping these cuts and accepting the somewhat larger systematic error on the flux.

Table 7 shows the signal acceptance, normalization acceptance, and apparent flux for the $K_L \to \pi^0\mu e$ analysis.

6 Results

Finally, Table 8 shows the expected backgrounds and the actual number of observed events for all three modes studied in this note. For the background in $K_L \to \pi^0\mu e$, I had added the K_{e4} background estimate to the estimate obtained from the $M_{\gamma\gamma}$ study, since the $M_{\gamma\gamma}$ sidebands would exclude any K_{e4} contribution.

Table 9 shows the final result for background, single event sensitivity (SES), and limits for each mode.

Cut variation	Variation 99 data	Variation summer 97	Variation winter 97
BA1 cut removed	4.1%	0.7%	3.4%
Extra track segs	3.5%	2.3%	2.7%
Extra DC hits	5.0%	4.6%	4.1%
Cluster energies > 3 GeV	0.3%	1.5%	2.0%
Track matching in magnet	1.8%	0.7%	1.4%
Total CsI energy	0.3%	1.5%	1.5%
Vertex χ^2	2.4%	0.7%	0.7 %
Total	8.0%	5.7%	6.7%

Table 6: Summary of systematic errors on the apparent flux due to cut variations from the normalization mode $K_L \to \pi^+\pi^-\pi^0$.

$K_L \to \pi^0\mu e$ data period	Signal Accep.	Normal. Accep.	Flux (10^{11})
99 data	3.95%	5.74%	3.40 $\pm$0.27
Summer 97	3.70%	5.70%	1.30 $\pm$0.074
Winter 97	4.11%	6.22%	1.47$\pm$0.098
Total flux			**6.17$\pm$ 0.30**

Table 7: Summary of acceptances and apparent flux for $K_L \to \pi^0\mu e$.

The final 90% CL limits were obtained by stepping through a range of branching ratio values which were known to include the final result. For each value of the branching ratio, I calculated the mean number of total (signal + background) events expected for that value of the BR. Then this value was allowed to vary according to the errors on the SES and the background. I generated a Poisson distribution at each value of the BR, then constructed an acceptance interval for that BR using the prescription of Feldman-Cousins. After stepping through a large range of BR values I arrive at the confidence belt constuction shown in figure 20. The limit is found by reading off the N=0 value, since in all cases here the number of observed events was zero.

Decay	expected in signal region	expected in blind region	obs. in signal region	obs. in blind region
$K_L \to \pi^0\pi^0\mu e$	0.44 $\pm$ 0.12	0.43 $\pm$ 0.10	0	0
$\pi^0 \to \mu e$	0.03 $\pm$ 0.015	0.03 $\pm$ 0.015	0	0
$K_L \to \pi^0\mu e$	0.66 $\pm$ 0.23	4.21$\pm$ 0.53	0	5

Table 8: Summary of expected backgrounds and observed events in the signal regions and the blind or control regoin for all three decay modes studied in this note. The background expectations for 99 and 97 datasets a have been added.

Decay	Expected bkgnd	Obs.	$SES^{-1}(10^{10})$	90% CL lim
$K_L \to \pi^0 \mu e$	0.66 ± 0.23	0	2.43 ± 0.125	7.56×10^{-11}
$K_L \to \pi^0 \pi^0 \mu e$	0.44±0.12	0	1.27 ± 0.051	1.59×10^{-10}
$\pi^0 \to \mu^{\pm} e^{\mp}$	0.03 ± 0.015	0	0.675 ± 0.027	3.59×10^{-10}

Table 9: Final results for backgrounds and BR limits

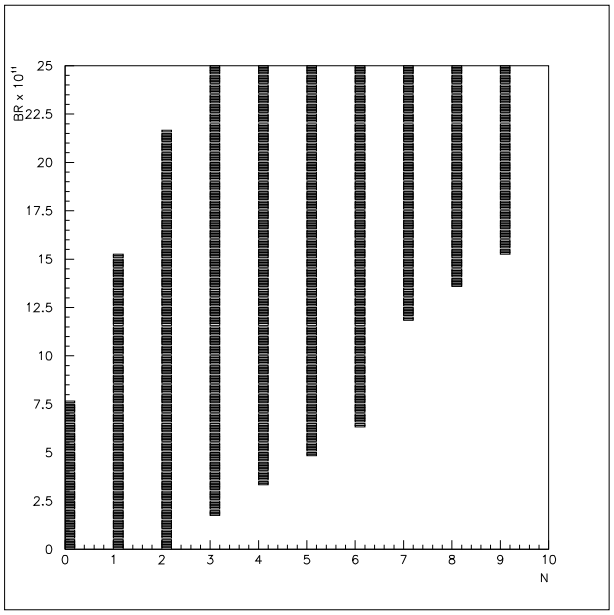

Figure 20: Example of a confidence belt construction, in this case for $K_L \to \pi^0 \mu e$.